THE OFFICE OF THE STATE CHIEF INFORMATION OFFICER
ENTERPRISE TECHNOLOGY STRATEGIES

North Carolina Statewide Technical Architecture

Domain White Paper

Network Architecture Technology Overview

STATEWIDE TECHNICAL ARCHITECTURE

Domain White Paper: Network Architecture Technology Overview

Initial Release Date:	August 1, 2003	Version:	1.0.0
Revision Approved Date:	Not Applicable		
Date of Last Review:	March 11, 2004	Version:	1.0.1
Date Retired:			
Architecture Interdependencies:			
Reviewer Notes: This is a move without modification from the old format to the new format as provided herein. A subsequent review and update will occur at a future, as yet to be determined, date. The Directory Services section was included and approved 3/7/2000. Published August 1, 2003 Reviewed and updated office title and copyright date. Added a hyperlink for the ETS email – March 11, 2004.			

Office of the State Chief Information Officer
Enterprise Technology Strategies
PO Box 17209
Raleigh, North Carolina 27699-7209
Telephone (919) 981-5510
ets@ncmail.net

Mission Statement

Network Architecture defines a common, uniform network infrastructure providing reliable and ubiquitous communication for the State's distributed information processing environment.

The Network Architecture specifies how information processing resources are interconnected, and documents the standards for protocols (for network access and communication), topology (design of how devices are connected together), and wiring (physical medium or wireless assignments).

The Network Architecture defines a unified, high-speed statewide network based on open systems standards. The biggest benefit to a statewide network solution is the ability to efficiently share information processing resources across the enterprise. Sharing resources is a common theme in all aspects of the Statewide Technical Architecture because economies of scale and efficiencies in operation result from collaborative approaches to technology. When agencies share common application services and data, they avoid duplicative efforts and costs. The key to successfully sharing these resources is a network connecting all state agencies together in a way that reduces redundancy.

A statewide telecommunications network must be strategically planned, strongly backed, and expertly managed. This network must:

- Utilize standard communication protocols.
- Sustain and support high capacity and high performance communication.
- Be scaleable, reliable, and extensible.
- Provide a variety of advanced telecommunications functions.
- Smoothly integrate with other private and public communication networks.

The computing industry has been going through an evolution. In the early years, a centralized processing environment would use a single-footprint large mainframe with a network of dumb terminals for access. Over time, as the capability and capacity of computers has improved; the trend has been to distribute the processing workload over multiple computers at multiple sites. The first step in this trend was the minicomputer environment where smaller versions of the mainframes provided localized networks for dumb terminals. Then, with the development of the personal computer, processing capability was distributed all the way to the desktop. Although such desktop capacity provided extreme flexibility and responsiveness, there were limitations. It was troublesome to share data. Users and business units

could not function in an interactive manner across the technology platform. To address these limitations, the concept of a local area network (LAN) evolved. A LAN allows multiple computers, printers, and peripherals to be interconnected and to function as a shared computing environment. Thus, the term computer has grown to imply more than a single personal computer or even mainframe. Instead the term computer may be used to refer to the entire collective group of networked computers, PC's, mainframes, workstations, and network servers, working together to provide a single information infrastructure. Within this new "computer", applications are distributed so the processing power is shared among several machines.

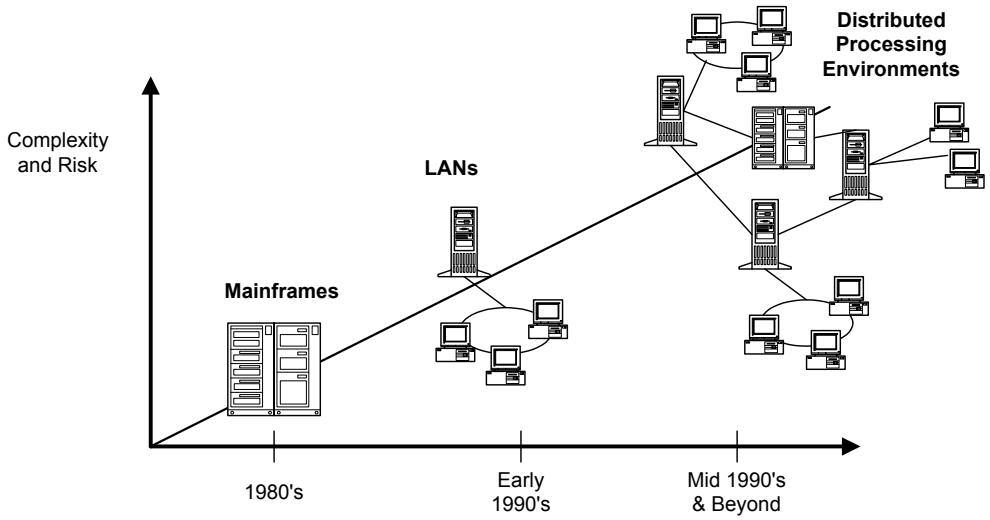

Figure 1. Changes in Network Environment.

The following types of networks are discussed in this chapter:

- ***Local Area Networks (LAN).*** A data communications system of multiple interconnected data terminals, computers, or devices confined to a limited geographic area consisting of a single building, a cluster of buildings, or a campus type of arrangement. The network does not use common-carrier circuits, although it may have gateways or bridges to other public or private networks.

- ***Wide Area Network (WAN).*** A data communications system that serves a large geographic area. WANs are often implemented using common-carrier provided lines. A WAN typically serves as a customized communication "backbone" that interconnects all of an organization's local networks with communications trunks designed to be appropriate for anticipated communication rates and volumes between nodes. The existence of a WAN permits the deployment of file, print, or application servers across the infrastructure to create centrally managed LANs where the close proximity of

components is no longer a requirement. The North Carolina Integrated Information Network (NCIIN) provides WAN functionality for state and local governments.

- ***Internet***. A *limitless* collection of interconnected LANs and standalone computers working in a cooperative manner under the standards and guidelines of the Internet Society.
- ***Intranet***. A limited collection of interconnected LANs and standalone computers. An intranet functions the same as the Internet, using the same user interfaces and file transfer protocols. The difference between an internet and an intranet is that an intranet provides connectivity between specific sites in order to create a pre-determined infrastructure for business units, customers, or designated participants. An intranet is often protected from outside access by a firewall. A firewall typically consists of a router with packet-screening ability that can block traffic between networks or specific host computers.

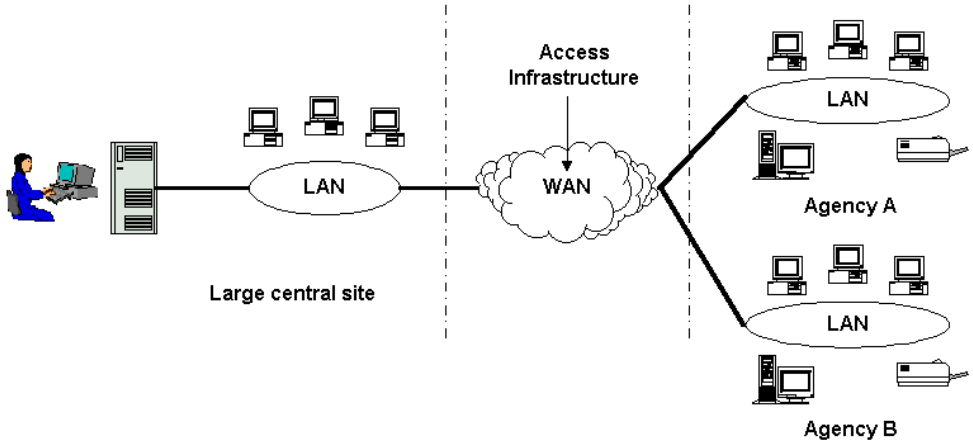

Figure 2. Statewide Network Architecture.

LANs support the needs of individual workgroups or individual agencies, but WANs support the cooperative and collaborative functions within the corporation or enterprise. Business requirements will necessitate using a variety of applications on the networks. However, the same need for variety does not exist within the network infrastructure. An uniform network architecture will enable LANs within the WAN to interoperate while allowing a broad platform on which to run applications as needed. Such interoperability requires cooperation at all agency levels and consistency in network components (e.g., wiring, hubs, servers, operating systems, and protocols), management practices, and services. The LAN and WAN sections of this chapter specify how this will be accomplished. The Network-centric Applications section explains how to design and implement business applications that effectively employ the network resources.

Note: Additional information about Wireless Networking, Inter/Intranet, Voice Services, and Video Network Services will be added to this Network Domain in future releases.

Local Area Network (LAN) Architecture

The invention of the microcomputer brought computer-processing power to the individual desktop. Users were able to have hands-on control of the accumulation, manipulation, and display of their own information. As such functionality became commonplace in the work area, users began to recognize the need to share information and resources with other users in their immediate area. Thus, local area networks (LANs) were developed to connect devices, such as standalone workstations and printers, in a limited geographic area such as a single building, a cluster of buildings, or a campus type arrangement. The initial LANs simply offered a means for users to share system resources, such as information and input or output devices. A network operating system (NOS) was used to accomplish this set of functions for the LANs. Over time, users began to require more functionality from these simple networks. Users wanted to expand the functions that could be performed on a LAN. Users also wanted to communicate with users and sites outside of their own work area. Thus, LANs were enhanced to offer support for a multitude of business applications. Application servers were added to the LANs in order to provide a multi-faceted, flexible environment. In addition, technology was augmented to allow communication between LANs through wide area networking techniques.

Technology Components

The following technical components are necessary for the successful implementations of LANs:

Topology

The way a network is physically wired refers to the network topology. There are three topologies used today; bus, ring, and star. (See Figure 3.)

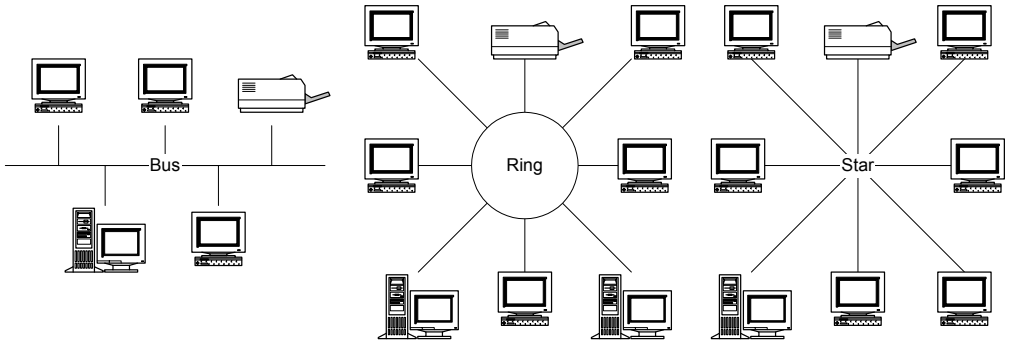

Figure 3. LAN Topologies

In a bus topology each device is connected to the network in a sequential manner so that, if the connection of one device on the LAN fails, the whole network fails. A ring topology connects devices in a closed loop. As with the bus topology, a problem with a single connection in a ring network will cause a failure from that point on the ring to the end. The star topology uses a central hub to which each network device is connected. Problems with a connection in a star network only affect that one device. Because each network device is attached individually, a star topology provides the capability to easily add and remove devices as necessary. Today's computing environments are dynamic infrastructures in which network configurations are constantly tuned, upgraded, and modified in order to meet changing demands and technology innovations. The star topology responds well to such demands.

Protocol

For two devices to communicate within a network they must both speak the same language or protocol. When devices do not use the same network protocol, they must use an interpretive device, or gateway, that translates the language one device uses into a language another device can understand. Coordinating the use of protocols across an enterprise WAN enables the organization to minimize the necessity for translations and, thereby, reduce support and capacity requirements.

The Institute of Electrical and Electronics Engineers (IEEE) is the organization primarily responsible for establishing standards for network protocols. The two most widely used standards are IEEE 802.5 (Token Ring) and IEEE 802.3 (Ethernet). Emerging high performance protocols include the Asynchronous Transfer Mode (ATM) and 100BaseT Fast Ethernet. These two offer multi-media communication technology that can handle telephony and video as well as conventional data.

The most widely accepted forms of Ethernet are 10BaseT Ethernet and 100BaseT Fast Ethernet. The number in the name stands for the signal speed in megabits per second (mbps). The Base means that devices on the network transmit using the

network's entire bandwidth (e.g., 10 mbps or 100 mbps). The T stands for twisted pair wiring. Used in a star topology, 10BaseT and 100BaseT are a reliable, scaleable, and maintainable choice. 100BaseT Fast Ethernet has the bandwidth necessary to support the needs of future voice and video requirements.

Cabling

The basic component of each network topology is the cabling. The options for cabling the network depend upon the particular requirements of LAN. Factors such as the distance between devices, volume of throughput, and network topology can determine what type of cabling is best suited for the LAN. When looking at the cabling of a network, the rule of thumb is that the technology used should have an expected useful life of approximately 15 years. The types of network cabling most commonly used are:

- ***Twisted pair***. Twisted pair cabling comes shielded (STP) and unshielded (UTP). STP cable is primarily used in Token-Ring environments. UTP supports almost all network applications such as voice, Token-Ring, Ethernet, and even Asynchronous Transfer Mode (ATM). A chosen Category 5 UTP should be certified for 100 mbps. Newer grades of UTP are being developed to support higher speed technologies such as ATM at 155 Mbps.
- ***Coaxial cable***. Coaxial cabling is generally used for video applications.
- ***Fiber optic***. Fiber optic cable uses light impulses instead of electrical impulses to transmit data from point A to point B. It can carry a signal further than copper cabling and can meet demands for higher bandwidth. It is often used in conjunction with other cabling to provide a network backbone between hubs and switches.

Hubs and Switches

In a typical Ethernet network each device is cabled to the LAN hub in a star topology. This hub contains one port for each device connected to it. Hubs act as the LAN "traffic cop" allowing streams of information traffic to flow between the ports in an orderly manner. In the event a port is busy the hub provides a buffer to hold the information until the port is freed up. The hub is an ideal point for network management due to its central location and because all network traffic flows through it.

Network switches provide the ability to break a network up into smaller sub-network segments. Switches are devices connecting segments of the network together. (See Figure 4.) Switches can be used in conjunction with or instead of hubs. They are used to improve LAN performance. Through the use of switching, network traffic is balanced across multiple segments thus reducing resource

contention and increasing throughput capacity. In addition, switching allows networks to assign increased speed or performance capability to particular segments in order to respond to heavy usage or application requirements.

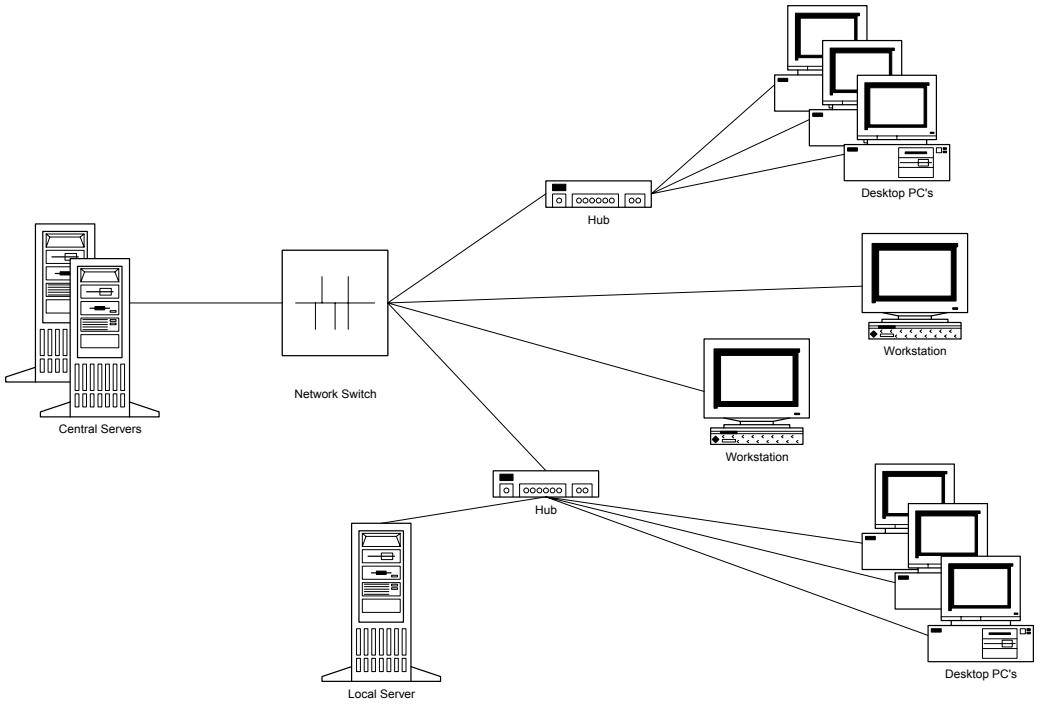

Figure 4. Switched Network

Wide Area Network (WAN) Architecture

A WAN is used to connect distributed network sites via private or public telecommunication lines. It typically serves as a customized communication "backbone" interconnecting all of an organization's local networks with communications trunks that are appropriate based on anticipated communication rates and volumes between nodes. The existence of a WAN permits the deployment of file, print, or application servers across the infrastructure to create remotely managed LANs where the close proximity of components is no longer a requirement.

The use of WAN technology may include both intranet and Internet access. With intranet access, an enterprise can provide connectivity between multiple network sites within the organization regardless of physical location. Although connectivity is limited to the specific LANs designated for this particular intranet, the enterprise does obtain the ability to improve its business processes across functional unit lines. Economies of scale and efficiencies of operation can be created by treating the entire enterprise's information infrastructure as a coordinated resource.

Resources can be shared, information can be synchronized, and support can be streamlined.

Access to the Internet must be obtained from an Internet Service Provider (ISP). With the addition of Internet access, the local network or the enterprise intranet obtains the ability to connect with other WANs and computer sites throughout the world. While this improves the enterprise's access to information and expanded customer bases, it also increases the enterprises need for security and improved management procedures.

The North Carolina Integrated Information Network (NCIIN) provides WAN functionality for state and local governments as well as public schools (k-12). Since the mid 1980's, the state has been planning, building, and managing a statewide-integrated digital WAN called the North Carolina Integrated Information Network (NCIIN). The NCIIN provides high speed, advanced function connectivity to meet public agencies' interoperability requirements. It is a telecommunications infrastructure that provides electronic access to services and supports the exchange of information between state agencies, no matter where they are located within the state. The telecommunications services supported by the NCIIN include:

- Data
- Voice
- Video
- Image

All NCIIN WAN services are provided to state agencies as a "turnkey" service offering by SIPS/STS. The STS staff provides planning and implementation services as well as operational and support services.

The state's WAN (NCIIN) architecture includes these facilities:

- ***Network Backbone Facilities***. Backbone facilities provide for the aggregation of an array of information, data, voice, video, and image services into a statewide transport mechanism. Compatibility with network access facilities is vital for connectivity. The planned future expansion of the network backbone facilities into regionalized network hubs will provide localization of traffic, improve performance and provide lower access costs per unit for transporting information. Migration to new technologies such as Synchronous Optical Network (SONET) will pave the way towards advanced, cost-effective high-speed connectivity.
- ***Network Access Facilities***. Access facilities provide a point of entry into the statewide network infrastructure. In today's environment, network access facilities are provided by a variety of technologies, including terrestrial fiber, hybrid coaxial wire, satellite, and wireless technologies. Access to the NCIIN

should be transparent to the other connectivity components and compatible with industry-established standards.

The state's strategy is to provide cost effective, ubiquitous service by leveraging its buying power. This strategy provides a variety of network access facilities, ensuring a range of services and efficient unit prices for clients.

Technology Components

The following technical components have been identified as necessary for the successful implementation of the statewide WAN.

Protocols

A communications protocol is a set of rules governing how computers exchange information with each other. Protocols were originally proprietary, for example IBM's SNA and Digital's DECnet. Today, the Internet uses protocols based on open standards. This permits connections between machines from many vendors.

Customer Premise Equipment (CPE)

Resources must be installed at the customer location to provide access from the local network to the WAN. This equipment includes:

- ***Router***. A router is a device that connects separate networks together. It forwards information packet via a determined best path through the WAN.
- ***Channel Service Unit (CSU).*** A digital interface device to connect end-user equipment to the local digital telephone loop.
- ***Digital Service Unit (DSU).*** A device used in digital transmission for connecting data transmission equipment (DTE) such as a router, to data communication equipment (DCE) or a service.

Carrier Services

These are comprised of various networking technologies offered by the telephone companies as a service. Services are typically broken into two categories: Switched (ATM, SMDS-DXI, Frame Relay and ISDN) and Non-Switched (DDS - Point-to-Point).

Internet Access

The Internet is a collection of networks with bridges or gateways between them. The protocol used by these networks is TCP/IP. Access to the Internet must be obtained from an Internet Service Provider (ISP). Connectivity to the ISP is mainly obtained in of the following ways:

1. ***Direct***. The local network is connected to a WAN with Internet services. Any computer on the local network can then access sites on the Internet via the network. This access is permanently available.
2. ***Dial up access***. A computer can use a modem (internal or external) to make a phone call to an ISP or to access another computer that already has direct Internet access. The dial up connection can be generated using the Point-to-Point Protocol (PPP) or Serial Line Internet Protocol (SLIP). This allows the local computer to become a node on the Internet, with an IP address, for the period of connection. The machine dialed up will either assign a specific IP number each time (server allocation, PPP and SLIP) or give a number currently not in use (dynamic allocation in PPP). This access is only available for the length of the phone call session. When the phone connection is terminated, the access is terminated and must be re-dialed to begin again.

Network-centric Applications

Designing and developing distributed client/server applications can be challenging because of the complications of a network. Network-related complications include:

- Moving data over a network is slower than moving the same data inside a single computer. The longer time required moving data between components of the application or between applications affects application performance. This time span is called *latency*.
- Not all network links operate at the same speed. This causes end-to-end application performance to vary, depending on where users and application components are located.
- Other applications will be using the network, too, which may reduce the network bandwidth available to your application.
- Network links and servers on the network will occasionally be down or unavailable. This can impact application performance and robustness.

While network issues should not affect the coding of application, application designers should consider the impact of their application on the network. The network connects the user interface with the business rules and the business rules with the data access code. (See Figure 5.) Application design decisions affect the flow of data over the network. The design decisions impact performance. This section describes design considerations for effective use of the network by applications.

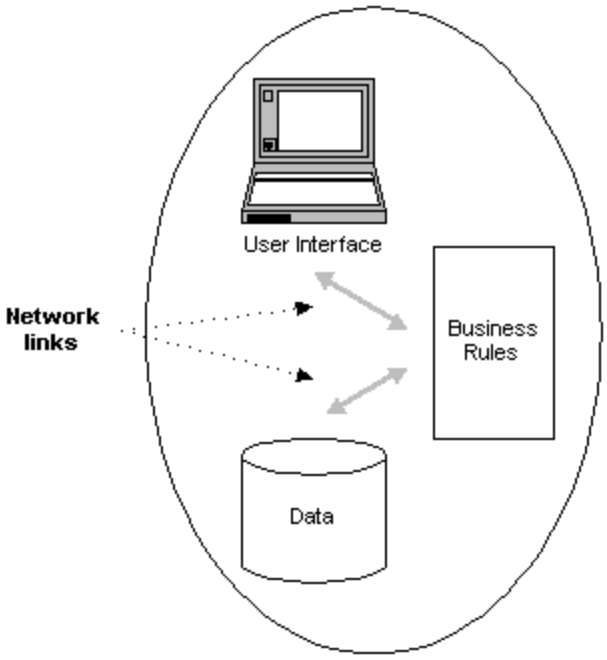

Figure 5. The network links between tiers of an application.

As applications communicate with other applications or with shared services, there are also application design decisions that affect the network. (See Figure 6.)

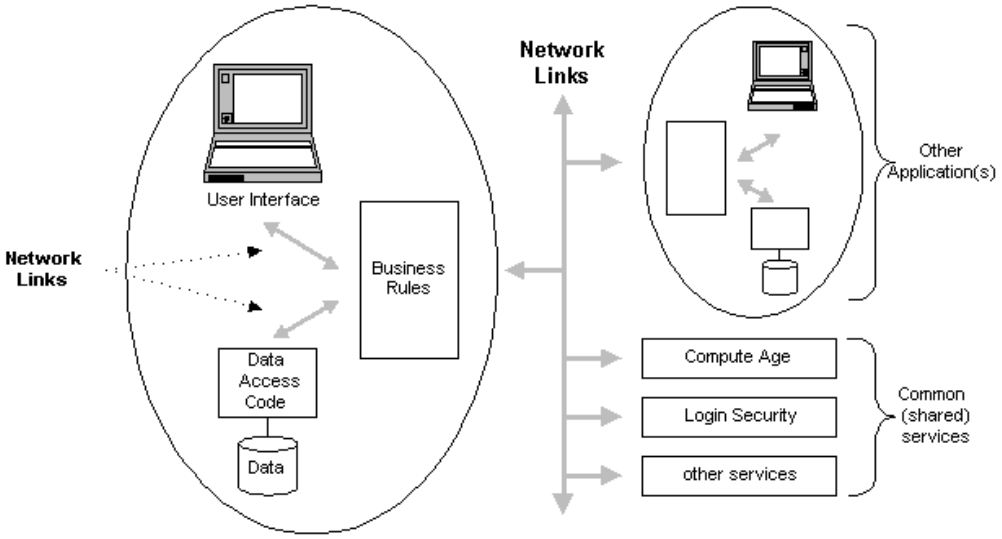

Figure 6. Network considerations impact communications outside the application.

There are 3 major areas of consideration for networked applications:

- *Infrastructure*: The planning and implementation of necessary network infrastructure must be driven by business requirements. Well-architected applications supporting most of the state's business can be supported by existing networks (LANs and WAN) Sometimes, business requirements will dictate the need for some out-of-the-ordinary network support. For example, if business requires high bandwidth (e.g., for video or to replicate large databases), or requires a wireless network (e.g., for remote users, such as county agents or EMS crews), the connectivity requirements must be understood before application design begins, so the infrastructure can be available when the application is deployed.
- *Application design*: Designers should make no assumptions about where application components will be deployed, or about the network bandwidth connecting the platforms on which it is deployed, unless business requirements dictate otherwise. If special networking is required, design with the knowledge that business-supporting infrastructure will be available. Otherwise, it is safer to assume a slow link will connect application tiers and to design accordingly.
- *Application deployment*: Deploy applications and applications components according to business needs. Except for "special cases," business needs are driven by the security and performance characteristics of each application within the context of the universe of applications supporting the business.

Directory Services

The business of the state is becoming more distributed. The state is developing closer electronic partnerships with businesses outside of state government, some employees are mobile users, some employees are working from their homes, and state services are being brought closer to the citizen electronically. This will require authentication services and a common enterprise repository of digital certificates that secures and supports E-commerce applications. Additionally, the state's technological resources must be available to users across the enterprise regardless of location or platform. To meet these goals, an enterprise directory services infrastructure must be in place. The state's technological resources and users are defined to the enterprise directory and appropriate access controls are applied. A directory is a natural place to provide security and it is the most important function it offers. It is the vault that contains the most trusted and critical components of an enterprise security strategy.

A directory service is a database that provides a mechanism to inventory, administer, and access resources in the network. These resources include users, groups of users, applications, data, printers, servers, and other physical devices throughout the network. A properly designed and implemented directory service

can present a central point for authentication (log-in) and a view of all available resources on the network. Additionally, this facilitates authorization, also known as access control, which determines the rights that are associated to a particular resource and enforces them. Directory services offer network users, administrators, and applications transparent access to all network resources and easy navigation of the network.

In the past directories were monolithic in design (See Figure 7), typically proprietary and embedded in each application. They also contained a limited amount of information, *i.e.* user ids, rights, and passwords. Although directories have recently become more distributed, many still have the limitations of a monolithic design – they are proprietary and offer little connectivity to other directories.

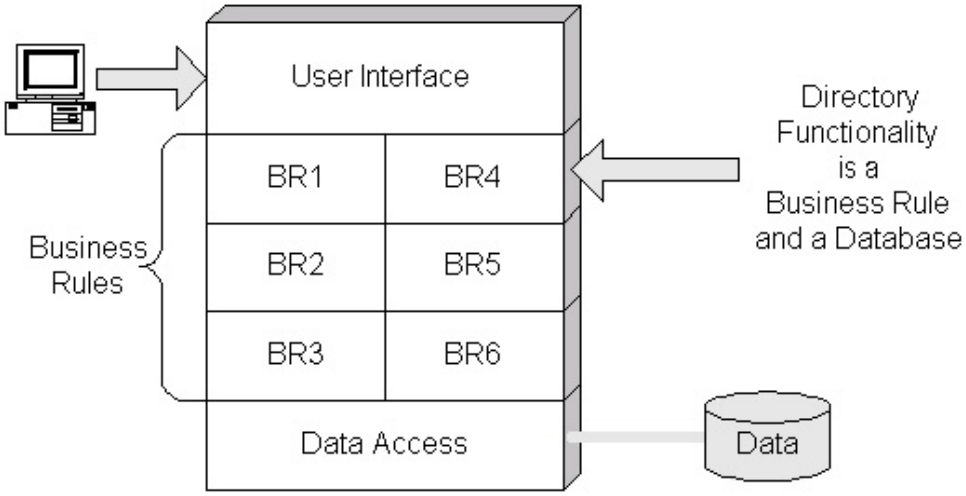

Figure 7. A Monolithic Application and Directory Architecture

Monolithic Directories

Directories are monolithic when the application and directory are tightly coupled. In essence the directory is built into, and becomes an integral part of, the application. The application uses its directory, implemented as a database or table, to look up a user's id and password during a login process. Many commercial off-the-shelf and older in-house developed applications have this architecture. This approach presents a number of problems, some of these are outlined below. *See the Application Architecture for additional information on Monolithic Applications.*

The drawbacks to directory services within a monolithic application (in addition to those outlined in the Application Architecture) include:

- *Scalability is limited by the application.* An enterprise directory strategy should have the ability to scale to millions of objects of multiple types. Monolithic directories within applications usually address only user accounts instead of

directory objects. These are usually not intended to support large numbers of users originating from a variety of platforms.

- *Integration with other applications or directories is difficult.* Interoperability and accessibility is critical to a distributed enterprise environment. Information must be available to users, other directories, and applications regardless of location or platform. Monolithic applications and the directories within are typically proprietary and offer no compliance to industry standards. This makes it difficult to integrate disparate environments. This also inhibits cross authentication, which provides a mechanism where an application can verify the authentication status of a user to an external source. This facilitates a decrease in redundant administration of accounts and requires fewer login requests of users.

- *Monolithic applications, and the directories within, have a single point of failure.* Since they are so tightly coupled, if any portion of the monolithic application or its directory becomes inaccessible, the whole application and directory becomes inaccessible.

Enterprise directories, based on industry standards, address the problems that exist in disparate proprietary monolithic directories. These solutions have been available for many years and the technology is mature.

Enterprise Directories

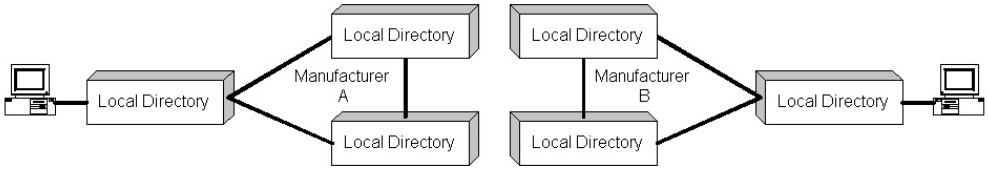

Figure 8. Enterprise Directories

An enterprise directory is a collection of local directories tied together to appear as one. This capability is usually limited to "like directories" from a single manufacturer. Enterprise directories must also have the capability to synchronize, replicate to each other, and have fault tolerant capabilities. When one of the directories in the enterprise becomes inaccessible, another will be available to service the requests.

Even within the directory products based on industry standards such as X.500, this is usually not possible unless the directories are purchased from the same manufacturer. This does not mean that they are not inter-operable. There are several options, such as meta directory, to synchronize dissimilar directories and offer some interoperability. However, that does not provide them with fault tolerant capabilities.

In an enterprise it is difficult to deploy one directory solution that all systems will use. Proprietary commercial off-the-shelf software, the lack of total interoperability in directory standards and products, and unique business requirements make this virtually impossible. However, in an enterprise one directory should be identified as the authoritative source for directory information and whenever possible utilize it for basic security services.

Meta Directory

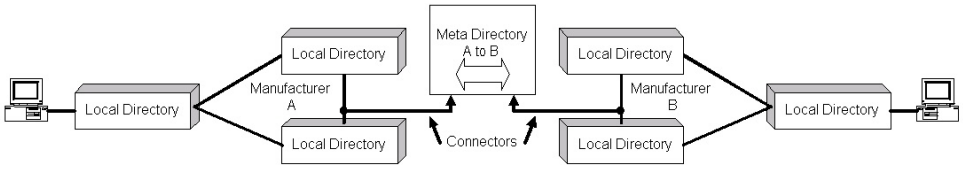

Figure 9. Meta Directory

Meta directories provide a mechanism to connect disparate directories by mapping directory schemata (See Figure 9.) The schema defines the characteristics of the directory such as object classes, access control information, as well as the relationships between the objects. Meta directories accomplish schema mapping through the use of connectors. This enables meta directories to connect many different directories to the meta directory without building one-to-one relationships (See Figure 10) to all of the directories. Additionally meta directories can maintain synchronization between the connected directories allowing a change to be automatically propagated to other directories. One directory, or the meta directory itself, must be identified as the authoritative source for directory information. Administration should be performed at the authoritative directory and propagated to the directory(s) that are connected. Many vendors offer this as an enterprise directory solution. However, the schema mapping can be very complex and difficult to maintain. Furthermore, any changes to the local directory attributes must be coordinated with the enterprise directory administrator and schema mappings changed. Changes are expected and commonplace, and a version change in the software can render a directory connection inaccessible. Meta directories can also be used to create unique user ids that can be spread to all connected directories. Today, meta directories are best implemented using vendor created and supported connectors, keeping the number of directories connected to a minimum.

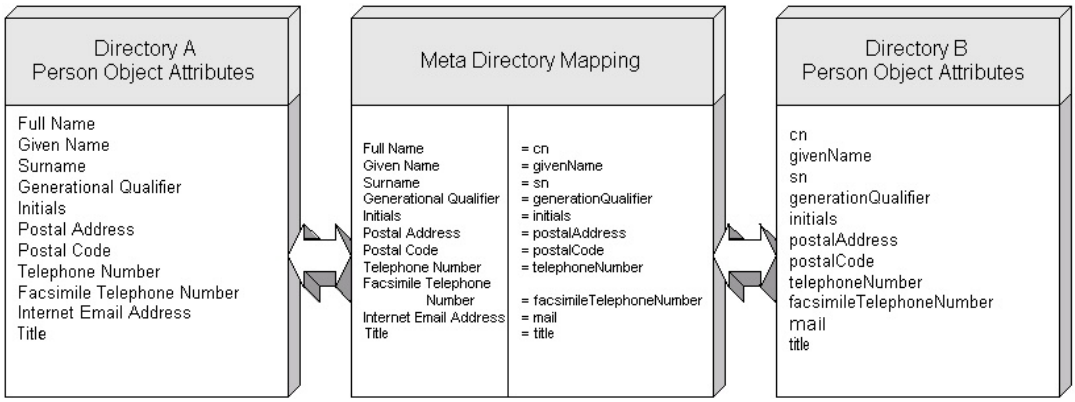

Figure 10. Simplified Meta Directory Mapping Concept

Directory Uses

Directories can enhance the functionality of other systems such as 1)-network operating systems, 2)-applications, and 3)-databases.

1)- Many Network Operating Systems (NOS) have their own integrated directory. Since a NOS is, by nature, a distributed environment with several servers and operating systems, NOS vendors have developed ways to connect and synchronize their directories. This offers users easy access to resources on all connected networks. The main purpose of NOS directories is to authenticate users and determine access rights to networks and resources.

2)- Applications require the ability to perform lookups in a directory to access data, authenticate users, and determine access rights. They can be directories enabled in order to obtain user authentication or authorization from an external source. Examples of applications that use directories include E-mail, Electronic Commerce, and Calendar/Scheduling. Applications that use an external directory service help reduce administrative burden by maintaining user account information in a separate location that is accessible to all applications. It also reduces the number of times a user is prompted to enter their login information. This is called a directory-enabled application.

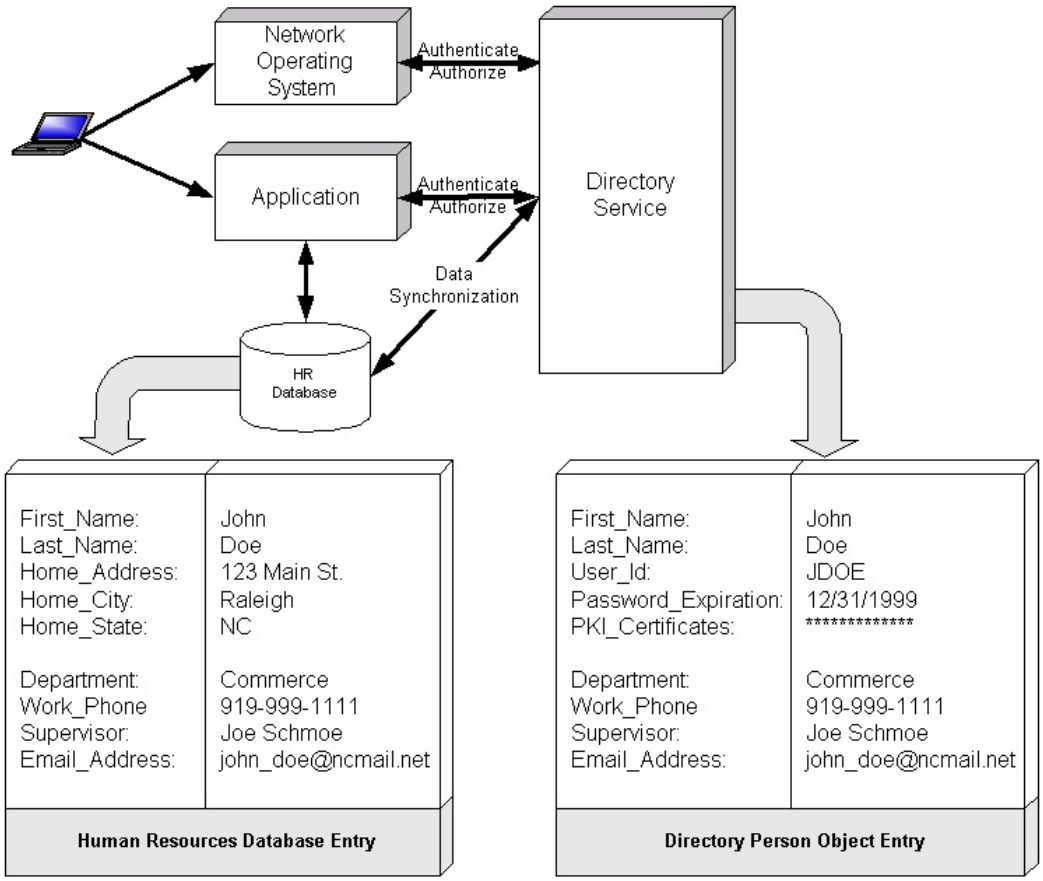

Figure 11. Directory Enabled Application

3)- Directory information, such as user information (*e.g.* phone number, address, etc.), is stored in the directory's database. Using database communications standards like ODBC, directories and application databases can compliment each other. This allows data to be gathered from multiple databases (*e.g.* a directory and an application database) to deliver a collection of information to an application. In this, data can be maintained in the most appropriate location and connected systems can access the information. For example, an application requiring human resources data could access the directory for certain user information and then access an application-specific database for any additional information needed (See Figure 11). Using the directory as a single point for common directory related information provides an efficient and effective database strategy and eliminate the problems caused by having duplicate data in multiple databases. The burden of redundant administration and the potential of unreliable and incorrect data would then be minimized.

Organizing Directory Entries

In order to implement operating systems and applications (along with their databases and directories) on an enterprise scale, disparate systems must have the

ability to communicate with each other. In response to larger more distributed networks, directory services have been improved to provide a means of organizing and managing the resources and communications between them. Today there are many standards and best practices that enable us to address these issues and implement an enterprise directory infrastructure.

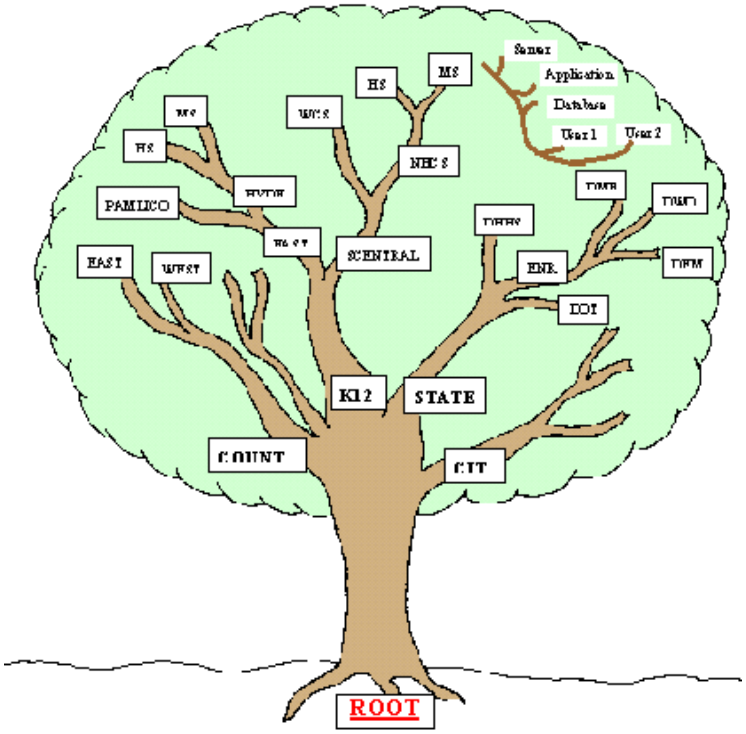

Figure 12. Conceptual Directory Tree

An enterprise directory infrastructure is made up of an expansive set of objects organized in a hierarchical tree (See Figure 12). Directory objects typically fall into two categories - physical and conceptual. Within these categories there are many object classes. Physical object classes are usually people, devices, or services (*e.g.* workstations, printers, applications, and data). Conceptual object classes could be any number of items including organizations, departments, or geographic object classes such as states, counties, or cities.

Directory Services Architectures Overview

All directory implementations have two general areas of functionality:

- *Security:* This includes authentication (log-in) and authorization (access controls) processes.

The directory is a natural place for security and is the most important function of the directory. It is the vault that contains the most trusted and critical components of an enterprise security strategy (*See the Security Architecture for more details*). The directory provides the information needed to secure network resources and applications from unauthorized usage. The identity of users is verified through credentials. The user can present credentials such as user name and password or digital certificate.

Since governmental business continues to be enhanced through the use of computer technology, protection of information and resources is a growing need. The authentication of individuals is essential to the security of the state's sensitive data.

An authentication service is used to identify and verify individuals, processes, and network resources. Authorization enhances security by regulating identified individuals access to protected resources. Authentication only ensures that an individual is really who they say they are; authorization specifies what they are able to do.

- *An administration facility:* Administrators must have a means of performing basic functions such as; adds, changes, and deletes to the resources defined in their local directory.
  Directories that exist in an enterprise must have additional features, some of which include:

- *Expanded administration capabilities:* Administration of the enterprise directory can be centralized, distributed, or a combination of both. In a distributed approach administrators can only make changes in their individual local directories. The local directories are then automatically synchronized to the enterprise directory. In a completely centralized approach, the single enterprise authority handles all changes and updates. A combination approach is, however, often the most effective and efficient. In such an approach, the responsibility for the consolidated enterprise directory is handled centrally while the responsibility for individual attributes is delegated to local workgroup administrators or even to the users themselves. For example, individual users might be allowed to manage certain elements of their own directory entry, *e.g.* their home telephone number or address. The changes can be queued for approval before committing them.

- *Scalability of the directory:* In an enterprise environment, as large as the state's, a directory must be able to scale to potentially millions of directory entries with little or no impact on performance.

- *Compliance to industry standards:* Interoperability with applications, databases, and other directories rely on industry standards as a means of communications.

- *Flexibility and adaptability:* Technology and business needs are constantly changing. Systems we install today must be able to adapt to the business needs

of tomorrow. They must be flexible enough to change with future technology. Most enterprise directory solutions today offer these benefits.

- *Fault tolerance capabilities:* One of the most valuable benefits of an enterprise directory is its tolerance to system faults. As stated earlier, if one directory in the enterprise becomes inaccessible, another directory will automatically service the requests.
- *Enhanced security options:* Cross-Authentication is a process where a user is authenticated across multiple directories. This reduces the number of sign-on requests a user is presented. An enterprise directory is critical to a Public Key Infrastructure strategy. It provides a central location for the storage and retrieval of digital certificates, revocation lists, etc.

Technology Components

An enterprise directory services system is comprised of the following components.

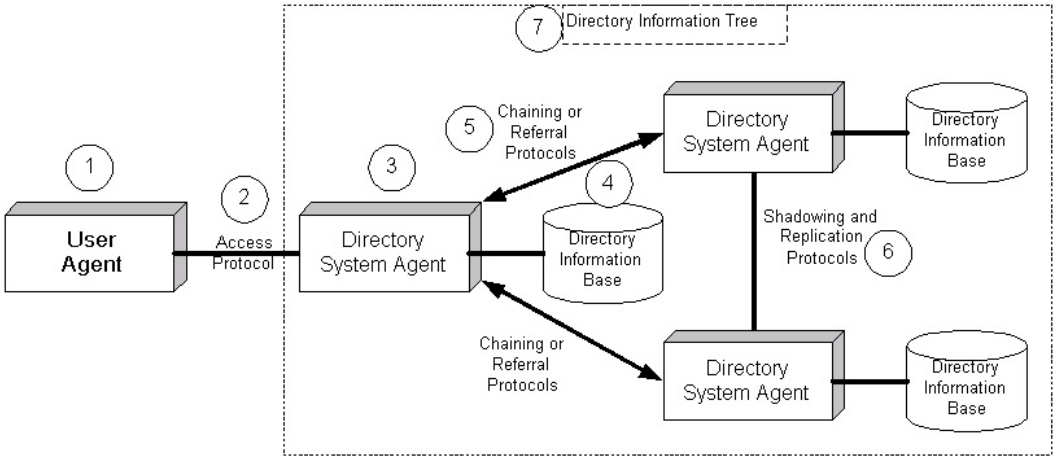

Figure 13. Directory Architecture

1 - User Agent

Access to the directory is through a user agent, also known as a user interface or a client. This does not refer only to a user accessing the directory, it includes whatever mechanism an application uses to communicate with the directory. This can be the traditional client application a user is presented during a login session. It can also be a service, either built-in or external, to an application that passes, for example, LDAP calls.

2 - Access Protocols

The user agent requires a protocol to communicate with the Directory System Agent (DSA) to request services (See Figure 13). User Agent services such as Read, Modify, Search, List, etc. request the DSA, using filters, to scan the Directory Information Base (DIB) and return the results. The user agent always initiates this communication. There are many access protocols. Some examples include Novell

Directory Access Protocol (NDAP), X.500 Directory Access Protocol (DAP), Open Database Connectivity (ODBC), and Java Naming and Directory Interface (JNDI). The Lightweight Directory Access Protocol (LDAP) has recently become the most popular and universally supported access protocol. As the name implies, LDAP is a lightweight front-end to many directories from many vendors. Initially LDAP was designed to increase performance and conserve bandwidth while accessing X.500 directories. Most directory and application vendors have chosen LDAP as an access protocol to communicate with their own directory or for their application to communicate with an external directory.

3 - Directory System Agent (DSA)
This is considered the directory server software. It performs all reads, writes, deletes, modifications, etc. to the directory information base (DIB) on behalf of the user agent via the access protocol.

4 - Directory Information Base (DIB)
The DIB is a database where directory information and objects are stored. Portions of this information can be distributed and replicated to other servers in the enterprise, in order to enhance performance and provide fault tolerance. Performance is improved by servicing requests for data from the DIB that is closest to the source of the request. Fault tolerance is achieved by replicating the data to multiple locations. If a DIB is unavailable, another server holding a replica of the information can then service the request. This strategy is found in X.500, NDS, and pure LDAP solutions. Although the techniques differ slightly, the concept is basically the same in each solution.

5 - Chaining or Referral Protocols
Chaining and referral protocols are the communications between DSAs, allowing a request to be passed from one DSA to another to obtain the information. Therefore, information in a directory can be accessed without needing to know the exact location of that specific piece of information. When a DSA receives a request, it queries the DIB. If the DIB does not contain the information requested, the DSA can then pass the request to another DSA to perform the lookup to its DIB and so on until a DSA can be found that has the requested information in its DIB. There are two techniques to accomplish this: chaining (found in X.500 directories using a Directory System Protocol - DSP) and referrals (found in Novell's NDS and Netscape's LDAP directory).

6 - Shadowing and Replication Protocols
To provide fault tolerance, multiple directories can be replicated to each other. This is usually only achievable with like directory types from the same manufacturer. Distributing copies of directory information to other servers allows another server to service requests when the destination server has become unavailable.

7 - Directory Information Tree (DIT)

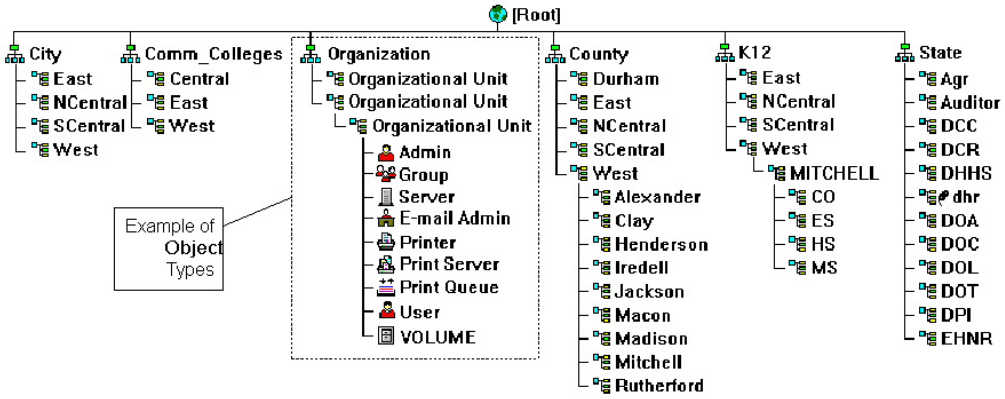

Figure 14. Directory Tree Hierarchy

The Directory Information Tree is a logical, hierarchical, representation of the enterprise directory as an inverted tree. Figure 14 is an example of a directory tree. This can be made up of millions of objects that, without a facility like this, would not normally be linked in any other way. This allows users and administrators to navigate through the tree without regard to the fact that they are spanning potentially hundreds of servers. Access controls determine what can be accessed or even seen in the directory tree.